

Who had an occupation? Changing boundaries in historical U.S. census data

Meyer, Peter B.

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Meyer, P. B. (2009). Who had an occupation? Changing boundaries in historical U.S. census data. *Historical Social Research*, 34(3), 149-167. <https://doi.org/10.12759/hsr.34.2009.3.149-167>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see:
<https://creativecommons.org/licenses/by/4.0>

Who had an occupation?

Changing Boundaries in Historical U.S. Census Data

*Peter B. Meyer**

Abstract: »Bei welchen Personen ist der Beruf bekannt? Wandelnde Kategoriengrenzen in der amerikanischen Volkszählung«. The original official purpose of the U.S. Census was to gather information to design political districts of approximately the same size. Increasingly Census data has been used for descriptive and social scientific purposes. This paper examines how the category of “occupation” has changed and looks at several issues which arise in comparing the present day workforce with the workforce in past decades. Changes in concepts, practices, and historical context have greatly affected how many persons were recorded as having occupations, especially for married women, American Indians, teenagers, and people who have ceased paid work.

Keywords: Longitudinal Analysis, Process-Generated Data, Social Bookkeeping Data, Public Administrative Data, Institutional Filters, Measurement, Census, Occupation.

1. Introduction

One way to trace social and economic changes is to track how the distribution of people in occupations has changed over time. The U.S. Census of Population has collected occupational information every ten years since 1850. In principle one could conduct a detailed analysis of the work force over time with this data, but there are a number of practical challenges in making comparisons. This paper describes some changes in the long run definitions of the work force and issues in extending the current measures of occupation into the past.

The concepts underlying the measures of the work force changed substantially between 1850 and 1940, after which they stabilized. This paper documents the changes in coverage and the conception of employment which make long-term comparisons difficult.

Since the 1930s, several basic census definitions have remained stable. A person is *employed* in a certain week if the person has a paid job or a job work-

* Address all communications to: Peter B. Meyer, U.S. Bureau of Labor Statistics, Office of Productivity and Technology, 2 Massachusetts Ave. NE, Washington DC, 20212, USA; e-mail: pbmeyer@econterms.net.

For their advice and encouragement I thank Margo Anderson, Nina Baur, Nele Bracke, J. David Hacker, Mindy Miller, Evan Roberts, Cindy Zoghi, and participants at seminars at the 2008 European Social Science History Conference and the RC33 2008 conference. Views and conclusions in this paper are those of the author and do not reflect views, policies, or official measurements of the U.S. Bureau of Labor Statistics.

ing at least fifteen hours a week in a family business producing goods for market. The employed person may not actually work at that job in that week if on sick leave, vacation, or other temporary leave. The *labor force* in a certain week is made up of the employed persons and the *unemployed* persons who are actively seeking employment during that week. These categories measure activity.

A person's occupation is a more complicated concept which has objective components and also may be integrated into the person's self-concept or social identity (Sobek 2006). For a person with a job, the occupation category is inferred by the Census Bureau based on the tasks or the job title. If the person has multiple jobs in a particular week, the person's occupation is the one which paid the most or at which the person spent the most hours. For a person who is not working, the occupation is self-declared and is expected to be related to the person's past work, training, or expected future work. There has been ambiguity and change over time in the likelihood that an occupation would be recorded for certain categories of people, such as a woman at home who is not currently working, or a man who has ceased working but perhaps not permanently. By modern definitions they are either unemployed, if seeking work, or not in the labor force.

Before 1940, census enumerators asked individuals for their occupations, and persons reporting an occupation were categorized as having *gainful employment*. This was the nearest substitute to the current labor force concept. The census did not carefully categorize the person's activity in a particular time period (e.g., working, looking for work, or neither) and so did not arrive at measures that neatly map into a labor force concept. There were also distinctions between who was recorded as having an occupation in the past and those who would be recorded as having an occupation now.

In the 1850 and 1860 censuses, the measures were quite different from those used at the present. In 1850 occupations were not recorded for women, and occupations were never recorded for slaves. This paper does not focus on that early period. Between 1870 and 1930, there were important differences in concept and application from current-day practice which this paper examines. Sections 2 and 3 review the historical decisions and definitions that led to the recorded occupation numbers. Section 4 then considers some specific subgroups for whom the patterns of occupation assignment changed or may not be obvious.

2. History of census occupations

The U.S. Constitution required the government to count persons by locality to help define representative political districts and proportional taxation across the states. As citizens moved and grew in number, the census data would be used to adjust the political districts electing the House of Representatives so they

would cover approximately equal numbers of citizens. A U.S. Census of Population has been therefore been conducted every ten years (“decennially”) since 1790. For the first century, these counts were collected by marshals, from the national law enforcement arm of the courts. The constitution required a separate count of the slave and free population. (Anderson Conk 1980: 8)

The Constitution did not call for collecting occupation or industry information, but it was repeatedly suggested, including by founder James Madison, that the census should collect such data for purposes of economic and military planning. The 1820 census include a question concerning how many persons, including slaves, worked in three general economic sectors, agriculture, commerce, or manufactures. By current conceptions, these are industry, not occupational, data. In 1840 respondents were asked the same question within these seven industry sectors: mining; agriculture; commerce; manufactures and trades; navigation of the oceans; navigation of canals, lakes, and rivers; and learned professions and engineers. Again it appears slaves were included (Hunt 1909). Starting in 1850, and ever since, respondents were asked more precisely for their profession, trade, or occupation.

In 1850, “profession, occupation, or trade” was requested of free males over 15. The results presented 323 occupations, summarized under ten general headings: commerce; trade; manufactures; mechanic arts and mining; agriculture; law, medicine, and divinity; other pursuits requiring education; government and civil service; domestic servants; and other occupations (Hunt 1909). The census director visited the statistical offices of several European countries in the early 1850s to discuss census practices. The 1860 census was very similar to 1850, but all free persons over 15, now including females, were asked an occupation. A list of 584 occupations was presented in the results (Hunt 1909). The census did not reorganize the reported occupations into a classification system, so similar occupations were sometimes reported separately or workers were grouped with others who did different kinds of work separately, and the Census Office did not make much use of the information. (Anderson Conk : 10) and Hunt 1909: 470) The Minnesota Population Center has applied the 1880 occupation classification to this data (see <http://ipums.org>).

In 1870, marshals continued to do the collection. The inquiry of occupation now went to all persons over 10 years of age, and the first classification system was applied to the responses. Histories of the occupation statistics (notably Anderson Conk 1980) emphasize the role which top census officials played in determining how jobs were reported. In 1870 and 1880, Francis Walker was in charge of the census. He tended to see the “sector” (industry) as important information to encode in the occupation classification.

Starting with the 1880 census, data collectors were no longer law enforcement officials, but rather temporary employees, who were sometimes hired through political patronage processes. The President appointed a Superintendent of the Census, and supervisors and enumerators were hired in lieu of the

marshals. The quality is said to have improved in each succeeding decade (Hunt: 1909).

Carroll Wright, an experienced statistician and government administrator, was in charge for 1890 and 1900. Hunt (1909) reported that quality improved with the 1890 and 1900 censuses. The Bureau redefined the occupational category system every decade, and expanded its detail on manufacturing jobs in the early 1900s when they were of growing interest.

There were more discussions of standardization across countries in the early 1890s. In 1893 the International Institute of Statistics defined a classification of occupations and in 1907 it sponsored a commission to prepare a multilingual glossary of the industries and occupations recorded in the censuses of the industrial countries.

The census was a large scale activity. Hunt reports there were more than 25 million copies of instructions and forms ("schedules") for the 1900 population census, several hundred supervisors, and 53,000 enumerators. The task was rushed.

Until this time, the Census Office infrastructure was temporary, and was recreated for a couple of years before and after each decade's canvass, and then shut down after the results were complete. After decades of advice and pressure on this point, in 1902 the U.S. Congress established a permanent Bureau of the Census to conduct the constitutionally required census and other surveys and analyses. This was expected to reduce loss of knowledge and skills between censuses and sustain a staff of professionals. The data collectors were now temporary civil service appointees. Other industrial countries made similar changes around this time.

A key analytical advance in 1910 was to ask two new questions along with the occupation question. One was what "industry" the employer was in. The other asked whether the respondent was an employee, or self employed, or an employer (Anderson Conk 1978). The existence of these related questions allowed the occupation question to focus on the worker's own tasks or function, not the source of an employer's revenues.

Alba Edwards was in charge of the occupational statistics from 1910 to 1940. He gradually reduced the use of industry information to classify occupations. He strongly preferred that occupation categories give information concerning the skills and intelligence of workers. He tended to elevate the importance of differences between jobs where workers "used their heads" versus those where workers "used their hands". In the absence of explicit data about occupational skills, the census was sometimes reduced to drawing inferences from the demographic characteristics of workers. He acknowledged that "certain specific occupations which technically are skilled occupations were classified as semiskilled [in census results] because the enumerators [found] so many children, young persons, and women as pursuing these occupations as to render

the occupations semiskilled, even though each of them did contain some skilled workers” (Edwards 1917).

From the point of view of current day concepts, Edwards improved the analytical precision of the occupation categories by helping separate industry concepts away from them. But in another sense he made the occupation concept more cloudy by including demographic or social-economic elements with structural or functional roles, for example by imprinting on them the idea that a woman could have a skilled occupation, but that an occupation made up mostly of women was not a skilled one. For a more thorough discussion see Anderson Conk (1978 and 1980).

During the Depression of the 1930s there was much discussion about how to make employment information more precise and timely, so that for example it would be possible to track, by a monthly survey, the effect of government programs. The gainful occupation concept was not precise enough to measure actual activity in a short period. After much research and discussion the government settled on the basic measurement categories of the labor force described earlier—the employed and the unemployed in a particular week. The existing categories of occupation and industry would continue but were decoupled from employment in a particular week. In the census of 1940 these concepts were applied, and subsequently a sample was surveyed monthly to track changes in unemployment. In the 1940s there were a number of redefinitions and clarifications (Durand 1948).

The 1950 category system is in wide use among academics. Matthew Sobek of the IPUMS.org project extended its reach to the entire span of 1850-2000, with jobs reported in the census matched as well as possible from the category system of the respondent’s year to the 1950 classification. Since 1950 the classification has been revised every decade, but the basic concepts of occupation, and who has one, have remained stable.

Table 1: Summary of census changes affecting occupation data

Year	Events
1790	Census begins, designed for political districting and taxation; most Indians not counted; slaves not individually reported.
1850	Free male respondents asked for their “Profession, occupation, or trade.” Occupations were recorded but not categorized in the census. IPUMS.org later put the responses into the 1880 occupation categories.
1850s	International conferences on occupation collection in censuses begin
1860	All free respondents asked for occupation; household head, usually male, is counted distinctively. Occupations were recorded but not classified substantively.
1870	Slave category disappears. Occupations classified into 338 categories.
1880	Data collectors now political appointees not law enforcement officials
1902	Census Office becomes a permanent civil service bureau. Data collectors were temporary employees.
1910	Industry and employer type are first asked separately from occupation, enabling better analytical separation of employer/employee concepts.
1924	A new law defined all American Indians to be citizens, so they were to be covered by the census.
1940	After much research and debate during the Great Depression, the census adopted de-gendered “labor force” definitions and concepts.
1960	Electronic “public use” samples first released from individual census data

The main classifications are in the next table. The respondent’s description of the work activity or job title was used by a marshal or Census employee to assign them into one of the categories.

Table 2: U.S. Census occupational classifications

Census year	Number of job names presented	The question asked, or other notes
1790-1840		No specific occupation question
1850	323	“Profession, occupation, or trade of each male person over 15 years of age” and “Number of slaves” (without further detail on their activities)
1860	584	“Profession, occupation, or trade of each person, male and female, over 15 years of age” and “Number of slaves”
1870	338	“Profession, occupation, or trade of each person, male or female”
1880	265	“Profession, occupation, or trade of each person, male or female” over age 10 and, of those, months unemployed during the census year. Separately, months at school. Full list of occupations is at http://usa.ipums.org/usa/volii/88occup.shtml
1890	218	“Profession, trade, or occupation” and, of those, months unemployed during the census year
1900	140	Many occupations were subcategorized by industry.
1910	215	Industry question now asked separately.
1920	224	Full list of occupations is at http://usa.ipums.org/usa/volii/92occup.shtml
1930	213	Full list of occupations is at http://usa.ipums.org/usa/volii/occ1930.shtml
1940	221	Full list of occupations is at http://usa.ipums.org/usa/volii/94occup.shtml
1950	287	Full list of occupations is at http://usa.ipums.org/usa-action/variableDescription.do?mnemonic=OCC1950
1960	296	Full list of occupations is at http://usa.ipums.org/usa/volii/96occup.shtml
1970	441	Full list of occupations is at http://usa.ipums.org/usa/volii/97occup.shtml
1980	504	“(a) What kind of work was this person doing? (at least two words) (b) What were this person’s most important activities or duties?” Full list of occupations is at http://usa.ipums.org/usa/volii/98occup.shtml
1990	504	Job list was similar to 1980. Full list of occupations is at http://usa.ipums.org/usa/volii/99occup.shtml
2000	510	Full list of occupations is at http://usa.ipums.org/usa/volii/00occup.shtml

Data Sources: For the count of 1850-60 categories, Hunt (1909). For 1870-1940 categories, Anderson Conk (1980: 23). For more recent categories, <http://usa.ipums.org>. For the phrasing of the question, Wright with Hunt (1900) and [ipums.org](http://usa.ipums.org).

3. Current census practices regarding occupations

In recent decades the answers of respondents to the occupation and industry questions in the decennial census have been encoded into a three-digit occupation code and a three-digit industry code, which are made available in the public-use micro samples. A monthly survey called the Current Population Survey, asks questions and follows procedures which are similar to the population census. This data is used to derive unemployment statistics. The following description of practices for the Current Population Survey (CPS) is similar to the practices for the decennial census. The CPS questions are available at <http://www.bls.gov/cps/bquestair.htm>, and the questions or instructions to the enumerators from past censuses are available at <http://usa.ipums.org>.

The respondent (who may be a family member or neighbor of the person being asked about) is asked approximately these questions in this order:

1) Class of employer:

- Is this person's employer a private for-profit company, a nonprofit, self-employed, federal government, state government, or local government?
- What is the name of the person's employer?

2) Industry:

- What kind of business or industry is it; what do they make or do?
- Is it mainly in manufacturing, retail trade, wholesale trade, or something else?

3) Occupation:

- What kind of work does this person do? (For example: plumber, typist, farmer)
- What are the person's usual activities or duties at the job? (For example: typing, keeping account books, filing, selling cars, operating printing press, laying brick).

The answers to these questions, four of which are open-ended text, along with the respondent's city, state, sex, age, and years of education are made available to specialized "coders" in the Census Bureau's National Processing Center in Jeffersonville, Indiana. I visited that office and interviewed some of the specialists.

The open-ended answers are in the respondent's handwriting, digitally scanned from the decennial census form, or were typed by the CPS interviewer onto a computer and then downloaded to the coder's computer. The coders follow carefully documented procedures to assign a three-digit industry code and an occupation code to the respondent. The open-ended answers are not made available in the public-use samples, and may or may not be available any more in any form.

The respondent may have given a job title, or identified tasks or activities at work. A normal problem is that the respondent has given too brief a description

of the tasks, or has exaggerated the importance or prestige of the job or tasks relative to the characterization in the classification system. A coder may not be able to assign occupation and industry codes based on the documented procedures, in which case the respondent's data is forwarded ("referred") electronically to a specialist called a "referralist." This occurred in 17% of cases in a large 1997-8 sample (Couper and Conrad 2001: 10). The referralist can match the employer name to a list of known employers called the Business Registry, and may research the employer or the words used in the answers to the "kinds of work" and the "activities or duties." The referralist may look up the employer on the Web, and also can refer to a number of books on occupations.

If the respondent's job is difficult to classify, there are residual occupation categories, usually with the text "not elsewhere classified" abbreviated in the title. For example, in the 1990 categories we see "Engineers, n.e.c.," "Managers and administrators, n.e.c.," "Therapists, n.e.c.," and others. This technique for inclusion helps some categories be precise while making it possible to include everybody.

4. Changing treatment of select populations

Some occupational categories have clear continuity over time, such as doctors, barbers, carpenters, laborers, or public officials. But much of the population was not categorized in an occupation, or was considered in an occupation in some decades but not others. A key principle was that a person's activity was almost always conceived of as an occupation if and only if the person received pay for their time. To clarify this idea through experience it is helpful to examine some of the sets of people who were sometimes thought of as marginally employed, outside the system of occupations, or outside the census for any reason.

4.1 American Indians or Native Americans

American Indians were not defined by the Constitution as citizens, and did not have representation in the U.S. Congress. Therefore, in the early decades they were not included in the census except for those who lived with American citizens. The sovereignty of North American Indian nations has been redefined over time. By 1890, there was an ethnic category in the census for "civilized Indians," meaning more or less those who had settled with Euro-Americans or African-Americans and assimilated to some extent. A 1924 law made all American Indians born in the territory of the U.S. citizens (Snipp 2006). There were special censuses of Indians on reservations in 1890, 1910, and 1930. These reported much larger numbers than in the other years. In other years, census enumerators did not search out Indians on reservations.

In 1960 and subsequently, the population was mostly self-enumerated, and the individual respondent could identify himself or herself as an American Indian. More people reported themselves as American Indians in 1960 and subsequently later than enumerators had counted in previous decades. (Alterman 1969: 300) Snipp (2006) reports that ethnic pride and a decline in stigma explains some of the increase. The numbers rose sharply again in 2000 as it was then possible for the respondent to report multiple racial backgrounds. 1.5% of the population reported some American Indian or Alaska Native background, and 0.9% reported no other racial background (Snipp 2006).

By the numbers reported in Snipp (2006: 767) and Alterman (1969: 293 and 300), many American Indians were not counted up through 1920. Overall totals are too few to generate the number of descendants now reported. A core reason is that the tribes were not all supposed to be counted in the census since they were sovereign nations or tribes, but it appears that through 1920 there were substantial undercounts even of the urbanized, Westernized population, because there were spikes and drops in the reported population.

In the 1900-1930 censuses, between 35% and 41% of the reported Native American population were classified with occupations, slightly less than the average for the rest of the population. One issue in counting occupations among them a century or more ago is that some subsisted on farming, hunting, gathering, or fishing for their families or tribes in a non-monetized economy and therefore might not have been counted as having employment according to a market concept even if they were making a living in their traditional way. Hunters and fishers are occupation categories in modern censuses but these are conceived as providers of goods or services to others, with payment expected.

4.2 Slaves

The Constitution required a count of free citizens and a count of slaves. The subject of slavery was very divisive politically, and after difficult bargaining the Constitution's framers agreed to count slaves as three-fifths of a person for purposes of a district's representation and taxation, though slaves could not vote and did not themselves pay taxes.

In 1850 census marshals asked slave owners how many slaves they had, and in 1860, asked owners for a list of the slaves' names. Because they were assets, owners were often able count them with precision. By 1870, after the Civil War, slavery had been abolished.

The census did not inquire about the occupations of slaves, nor was „slave“ defined as an occupation. Most of them were farm workers. Some slaves were specialized and skilled, and their work is the subject of a literature. One can see examples of their jobs as a tourist at George Washington's plantation, where among the slaves there were metalworking and leather specialists, clothing makers, food processors, housekeepers, and farm workers.

It has been estimated that 90% of the slaves over age 10 worked. If so, then on the order of 2.5 million slaves in 1850 had occupations and by 1860 perhaps 3.2 million did. These are very rough estimates. Economic historians of the period find that the ex-slaves worked significantly less once their time was their own after the end of slavery (Ransom and Sutch 1977, Weiss 1992, and Carter 2006: 19-20).

4.3 Adult women

Wives and mothers who were not employed outside the home were often not counted as having occupations in the period 1870-1930, even if they took in work such as sewing or child care at home, or had recently worked. Women in general were often not counted as having occupations because they were expected to be dependents, or thought of themselves in this way. There is a substantial literature on this topic.

Census administrators struggled with this issue and wrote slightly different instructions to enumerators over the decades. These are listed and discussed in Roberts (2007: 30-52). One recurring problem was that an enumerator might report a woman's occupation as "keeping house" or as "housekeeper." The usual rule was that the first activity was on behalf of her own family or household, whereas the second was a paid job. But it was found that this distinction was not perfectly well understood or maintained and that variations could have a large effect on the final numbers of employed persons (Roberts 2007 and Carter and Sutch 1996).

A number of researchers have tried to assign occupations to some of the women for whom no occupation was recorded. Goldin (1990) and Bose (2001) have for example matched married women in certain past censuses to likely occupations. In houses with boarders, they impute that the woman was a boarding-house keeper, and in farm houses they impute that the woman was a farmer. Goldin (1990) found that only 4.6% of married women were reported as having occupations in 1890 but estimates that 14.5 % were in the labor force. Bose (2001) evaluated how then-conventional roles for women affected the way 1900 census data reported female-headed households and formally-unemployed housewives. While the 1900 census reported that 22.5% of women had occupations and were employed, Bose (2001) imputes others to have been managing boarders, working on family farms or otherwise employed and arrived at a total of 46.4% of women aged 15-64 as employed with an occupation. Thus using present-day definitions, twice as many women were working in 1900 than were reported in the 1900 census. They would be recorded as having an occupation based on the same data in a recent census.

Carter and Sutch (1996) reported an important discovery, that according to the returns from the census enumerators of 1880, over 40% more women were employed than were reported in the official census totals. They identify differ-

ences in housekeeping versus keeping house as part of the source of the difference, but the data seem to have been adjusted for other reasons. It is possible that the same differences exist for 1870 and 1890 or other decades also. For 1890 this cannot be confirmed because the original 1890 reports on individuals were destroyed in a fire. The individual census responses for this period have not all been computerized and those that have are not all incorporated into the literature discussed above.

Roberts (2007: 4 and 390) concludes, after examining the census data on individuals and the literature of critiques and suggested corrections to this data, that the proportion of married women who worked declined slowly from 1860 to reach a low point between 1900 and 1920. Then more and more married women entered wage work until 1990, when this growth leveled off.

4.4 Can't find, homeless, refused, or travelling

If census enumerators cannot reach any person in the household, they can accept a secondary report on that household. If the person has no identified home but generally resides in a locality, this person is supposed to be included, and some homeless persons do have occupations. A person who is traveling should be recorded in the district where the person lives, not where they temporarily travel or work. This last criterion makes the U.S. census by definition a *de jure* not a *de facto* census. Thorvaldsen (2006) compares censuses internationally regarding their handling of this issue.

4.5 Children and students

There is sometimes a lower bound on persons for whom an occupation can be recorded. Based on the instructions to the enumerators, the lower bounds on age were specified to be 15 years of age in 1850 and 1860, 10 years in 1880 and 1900, and 14 years of age in 1940-1960. In other years there did not appear to be any restriction at the time of data collection. In recent years the lower bound is 16. These rule changes do not much affect the total number of persons with a recorded occupation because not many children below the minimum age for a given year are actually working. However the counts for particular occupations such as unpaid farm laborer and domestic servant could be affected meaningfully by the age changes (Sobek 2006: 36).

Being a student is not an occupation but in some years has been recorded in lieu of occupation. Now, a youth or student with a paying part-time job or unpaid family work taking 15 hours or more per week is to be recorded as having an occupation. Before 1930 there was not always such a firm quantitative criterion and so there was some vagueness or room for the respondent or enumerator to exercise judgment about whether the youth had an occupation. In general, scholars of this topic think that youth employment was underreported (Carter 2006: 16).

4.6 Retired, unemployed, or not in the labor force

Since 1940, a person who is not working and does not intend to return to work is categorized as not in the labor force, and also does not have an occupation (except when retired is recorded as an occupational classification). The delineation has become sharper over time. Since 1970 a respondent who does not have a job, but had one less than five years ago and expects to work again, may respond with an occupation which is the customary job or a previous job.

Moen (1994) summarizes a literature on the question of whether occupation data were collected on unemployed older men between 1890 and 1910, even when the persons were really retired permanently. In modern data, there is a retired category in lieu of another occupation. The census instructions over this period became clearer and clearer in their attempts to get enumerators not to record occupations for retired men, but they did not have a sharp category for this.

Comparisons of labor force estimates of various groups in 1940 with the earlier time series on gainful employment shows that although for most groups they were similar, older men were likely to report an occupation even if they were no longer in the labor force. Some of the older men reporting an occupation were actually more or less retired. This is the only substantial category of persons reporting occupations that would be smaller using the definitions now than using the definitions in use at the time. A smaller countervailing effect is that some of the men who were unemployed but planned to return to work apparently did not report an occupation. The problem in making an estimate may be partly that the respondents did not feel committed to being in a state of retirement, or not. A stated occupation may be an identity representing the current work one is doing, the work one has done, the kind of work one is searching for, or the kind of work one intends to do but is not searching for right now. These can be vague inchoate concepts the mind of the respondent so the count of persons with an occupation, both in the past and now, are not always as sharply defined as the current counts of the unemployed. For example, an elderly man with a tiny pension who sporadically searches for part-time employment might or might not think of himself as having a regular occupation with regular tasks, and in 1900 the distinction was less clear than now.

Estimates of how many older persons were retired in the modern sense before the 1940 vary greatly. The statistical bureau did not have a sharp technical distinction between the temporary unemployed and the permanently retired, and at the same time the persons being asked to define their occupation or status may not have had a sharp distinction in mind. Ransom and Sutch (1989) have made estimates of how many were retired around 1900. Moen (1994) takes the view that it is not clear how to distinguish the men who were retired from those who were unemployed and planned to work in their reported occupations.

Standardized public concepts of retirement and unemployment were firmly established in formal legislation on social security and unemployment compensation in the 1930s. Statistical practices by the mid-1940s established monthly measures of the unemployed category and therefore distinguished them more sharply from the permanently retired, although some ambiguity remains.

4.7 Non-citizens and border-crossers

The census is a survey of people who live in the United States. Likewise, persons crossing borders between home and work, legally or otherwise, are counted in the census and their occupations recorded if they live in the U.S. and not if they live in another country. (These numbers can be large. It has been estimated recently that more than ten percent of Mexico's population lives in the U.S.). For census measurement of "occupation" it does not matter whether the employer or the workplace is in the United States, although for the official redistricting purpose, only citizens count.

The United States expanded to include Alaska and Hawaii in 1867, which later became states, and to include Puerto Rico, Guam, the Virgin Islands and other territories in 1898. After they became U.S. territories, census data covered the people in these areas, sometimes in special reports. The occupation concepts and categories were the same as the rest of the U.S..

4.8 Military

All the persons in any armed forces are employed, and in most decades their occupations were recorded in a few special military-only occupation categories. So no one from the military was counted in the category of "doctor" or "manager" although these terms might describe their tasks.

For the 1990 census, there was a change in practice so that some military persons were categorized more closely by their work activity. For example, in the 1990 data, a physician in the military would be in the occupation "physician". That change brought those people in alignment with the usual principle that the person's work tasks or activities, not the employer's attributes, determines the occupation category. But in 2000, the census returned to a classification system in which anyone employed by the military is in a military-only occupation category.

4.9 The institutionalized

These are persons who on census day are hospitalized, imprisoned, in a government or charity shelter for safety, or in special institutions for the aged or disabled or mentally incapacitated. These persons count in the census, and can be reported to have an occupation but only rarely do. The rules about counting them do not seem to have changed significantly over the decades.

4.10 Illegally working, doing illegal work, or evading taxes

The census is supposed to include such persons and count their occupations. Individual data from the census is not made available to law enforcement officials. A basic principle is that the census data are to be an accurate count of persons, reporting the respondent's world to the degree the categories allow. Therefore a person employed in criminal activities, or not paying taxes on employment may simply report their activity as an occupation to the census. Gambling and prostitution are legal in a few places in the U.S. In principle, employment in these is reported as an occupation.

4.11 Volunteers, hobbyists, and persons at leisure

A person devoted to volunteering, for example giving tours at a museum without pay or serving a religious institution or political party without pay, is not employed, therefore these activities have not been counted as census occupations. A person delving into a hobby like stamp collecting is not counted as having an occupation. There is however a self-employed category, and a person trying without pay to invent something that will be valuable in the future may probably report himself as self-employed as an inventor, and perhaps this would match an occupation. The rules on counting such activity as an occupation do not appear to have changed over time.

4.12 Apprentices

There were many apprentice occupation categories until 1960. If payment was explicit, then an apprentice had an occupation. If there was no payment in the present but rather the prospect of a job in the future, it seems that this would be recorded as a student role, not an occupation.

4.13 Households

In the 1850-1860 censuses households were recorded as having an occupation, which was the principal occupation of the household head. Since then, the occupation concept is closely attached to individuals, and never attached to households.

4.14 Undercounts

By construction, census procedures normally conclude with an undercount, since procedurally the enumerators must identify each person before recording data about that person, and there is no compensating mechanism to estimate or impute data from those who are never located. People are more likely to be missed by the census if they are disconnected from jobs, families and stable homes. By some estimates blacks have been undercounted by over 10% in the

U.S. censuses of 1870, 1920, 1940, 1950, and 1960 (Coale 1955 and Alterman 1969: 275-280). The fraction of the population that has been counted has risen over time (Thorvaldsen 2006). Therefore over the decades the census is likely to have missed a declining fraction of the population who would report having an occupation.

5. Translation and standardization over time

Several efforts have been made to enable a researcher to map or translate an occupation in one census into a category system that is comparable to occupations recorded in other decades. The Minnesota Population Project's IPUMS project web site at <http://ipums.org> makes such category systems available.

Translating occupations into other category systems is helpful for comparisons, such as the efforts to infer the effect of unionization, licensing, or technological change. The "not elsewhere classified" occupation categories make it feasible to extend a classification system to cover multiple decades since they can incorporate occupations whose technologies or roles were not in use in both time periods. For example, when encoding 2000 data to the 1950 category system, there is no very good match for computer network administrators or computer support specialists. An n.e.c. category may be the best place to sweep them. This helps systematically address the goal of comparing work forces over time, but these particular categories include qualitatively very different kinds of work. Comparisons of data in these categories over time mix changes in a population over time with changes in the meaning of the category.

Large portions of the population are outside the occupational categories, and as discussed above the limits have been drawn differently over time. Whole categories of people have shifted from having no identified occupation, to having one. Specifically, slaves were completely outside the occupation category system. Most Indians were not in the mid-1800s citizens in the census. By one estimate, 46% of adult women were working in 1900 if we use current definitions, but only 23% were by the definitions in the 1900 census. Going in the other direction, retired people now are clearly defined as outside the occupation categories, but were not in the decades around 1900. Overall, much of the population has transitioned or been redefined into the officially measured labor force as financial and economic interaction has become predominant, over and above ethnic, racial, national, and gender categories. A variety of steps have been taken to impute occupation to people in the census for whom there is no occupation information, although there is no unified agreed-on alternative data set in which they are all taken into account.

6. Conclusion

There is an objective definition to whether a specific individual is in the U.S. labor force – either the person has employment, or the person made a specific effort to seek work in a specified week. Occupation is a more amorphous concept, which can reflect elements of a person's intended role and self-image. For example, a person who is not currently employed has some room to specify an occupation based on past work, past training, or future intent.

In 1870 and after, millions of workers were either not counted, or their occupations were not recorded. A line of research has made careful estimates of how many women, especially married women, had jobs but were not recorded as having occupations because the respondent or the enumerator did not think of the woman as having an occupation. Many American Indians were not counted because they were in sovereign states, or were not perceived as having occupations in the current sense. And many older or disabled persons who might not expect to work again were recorded as having an occupation because it was part of their ongoing identity, although now they might be categorized as retired.

The rules and system evolved toward defining more sharply each person's role in an imagined national economic order. These roles were recognized and regularized in the census as jobs and occupations chosen by the individual and representing a commitment to employment, rather than being tied to ad hoc activities, status, or family relationships. For example, over the decades, wives who worked on farms were increasingly likely to be recorded as a worker as well as a wife. The category of slave (which was not chosen) disappeared, and American Indians were brought into the regular category scheme. The census over time clustered persons into predefined occupation categories so that unusual cases would be standardized.

When a person did not have an occupation, the alternative categories became regularized so that each group could be counted. The definition of a child as working or not working became formalized. The individualized non-occupation categories of retiree and student became more formalized and grew. In each case the broader world, outside the census, evolved toward defining persons as autonomous but expected or obliged to choose a job or an occupation, and the employer as a member of a separately conceived industry concept.

In framing the data in this way, the Bureau of the Census was also responding to its users, supplying data which described the state of the national economy in an ever more timely and precise way. The original institutional role of the census did not require such a change. Instead, the census responded to national needs for information, as increasingly suggested and emphasized in social scientific descriptions of aggregate activity.

References

- Alterman, Hyman (1969): *Counting People: the Census in History*. New York: Harcourt, Brace & World.
- Anderson Conk, Margo (1978): Occupational Classification in the United States Census: 1870-1940. In: *Journal of Interdisciplinary History* IX. 111-130.
- Anderson Conk, Margo (1980): *The US Census and Labor Force Change: a history of occupation statistics, 1870-1940*. Yale University Press.
- Anderson, Margo (1988): *The American Census: A Social History*. New Haven: Yale University Press.
- Bose, Christine E. (2001): *Women in 1900*. Philadelphia: Temple University Press.
- Carter, Susan B. / Sutch, Richard (1996): Fixing the Facts. In: *Historical Methods* 29. 5-24.
- Carter, Susan B. (2006): Labor. In: Carter, Susan B. / Gartner, Scott Sigmund / Haines, Michael R. / Olmstead, Alan L. / Sutch, Richard / Wright, Gavin (Eds.) (2006) *Historical Statistics of the United States, Millennial Edition, volume 2*. Cambridge University Press. 1-35.
- Coale, Ansley (1955): The Population of the United States in 1950 Classified by Age, Sex, and Color – A Revision of Census Figures. *Journal of the American Statistical Association* 50. 16-54.
- Couper, Mick P. / Conrad, Frederick G. (2001): *Analysis of Occupation Coding in the Current Population Survey*. Report for Bureau of Labor Statistics.
- Durand, John (1948): *The Labor Force in the United States, 1890-1960*. New York: Gordon and Breach.
- Edwards, Alba M. (1917): Social-Economic Groups of the United States. In: *Journal of the American Statistical Association* XV. 645-653.
- Goldin, Claudia (1990): *Understanding the Gender Gap: An Economic History of American Women*. New York: Oxford University Press.
- Hauser, Philip M. (1949): The Labor Force and Gainful Workers – concept, measurement, and comparability. In: *The American Journal of Sociology* 54. 338-355.
- Hunt, William C. (1909): The Federal Census of Occupations. In: *Journal of the American Statistical Association* 11. 467-485.
- Moen, Jon R. (1994): The Unemployment and Retirement of Older Men: further evidence from the 1900 and 1910 Censuses. In: *Historical Methods* 27. 40-46.
- Ransom, Roger L. / Sutch, Richard (1977/2001): *One Kind of Freedom: The Economic Consequences of Emancipation*. Cambridge University Press.
- Ransom, Roger L. / Sutch, Richard (1989): The Trend in the Rate of Labor Force participation of Older Men, 1870-1930: A Reply to Moen. In: *Journal of Economic History* 49. 170-183.
- Roberts, Evan (2007): *Her Real Sphere? Married Women's Labor Force Participation in the United States, 1860-1940*. Dissertation, University of Minnesota Department of History.
- Snipp, C. Matthew (2006): American Indians. In: Carter, Susan B. / Gartner, Scott Sigmund / Haines, Michael R. / Olmstead, Alan L. / Sutch, Richard / Wright, Gavin (Eds.) (2006) *Historical Statistics of the United States, Millennial Edition, volume 1*. Cambridge University Press. 715-721.
- Sobek, Matthew (2006): Occupations. In: Carter, Susan B. / Gartner, Scott Sigmund / Haines, Michael R. / Olmstead, Alan L. / Sutch, Richard / Wright, Gavin (Eds.)

- (2006) Historical Statistics of the United States, Millennial Edition, volume 2. Cambridge University Press. 35-40.
- Thorvaldsen, Gunnar (2006): Away on Census Day: Enumerating the Temporarily Present or Absent. In: Historical Methods 39. 82-96.
- U.S. Bureau of the Census (1963): U.S. Census of Population: 1960. Subject Reports. Occupational Characteristics. Final Report PC(2)-7A. Washington, DC: U.S. Government Printing Office.
- Weiss, Thomas (1992): U.S. Labor Force Estimates and Economic Growth, 1800-1860. In: R. Gallman, J. Wallis (Eds), American Economic Growth and Standards of Living Before the Civil War. Chicago: University of Chicago Press. 19-78.
- Wright, Carroll D. / William C. Hunt (1900). The History and Growth of the United States Census. Washington, DC: U.S. Government Printing Office.